

DOI:10.19344/j.cnki.issn1671-5276.2025.06.003

基于 VMD-CWT 和 Swin Transformer 的滚动轴承故障诊断方法

曾信凌, 龙江, 魏友, 吴云飞

(中国民用航空飞行学院 航空工程学院, 四川 广汉 618300)

摘 要:针对滚动轴承故障信号存在噪声干扰且故障特征提取不精确的问题,提出一种基于变分模态分解(VMD)、连续小波变换(CWT)和 Swin Transformer 网络相结合的滚动轴承智能故障诊断方法。利用变分模态分解对信号进行降噪,通过 CWT 将重构后的信号转换为时频图;以二维特征图像作为输入训练 Swin Transformer 模型,实现滚动轴承的智能故障诊断。试验结果表明:VMD-CWT 结合 Swin Transformer 网络的方法具有更高的故障诊断精度,实测数据中测试集准确率高达 99.79%。

关键词:滚动轴承;变分模态分解;连续小波变换;Swin Transformer;故障诊断

中图分类号:TH133.33 **文献标志码:**A **文章编号:**1671-5276(2025)06-0018-06

Fault Diagnosis Method for Rolling Bearings Based on VMD-CWT and Swin Transformer

ZENG Xinling, LONG Jiang, WEI You, WU Yunfei

(School of Aeronautical Engineering, CAFUC, Guanghan 618300, China)

Abstract:In response to the noise interference and inaccurate fault feature extraction in rolling bearing fault signals, an intelligent fault diagnosis method for rolling bearings combining variational mode decomposition (VMD), continuous wavelet transform (CWT), and the Swin Transformer network is proposed. The signal is denoised using variational mode decomposition, and the reconstructed signal is converted into a time-frequency map using CWT. The two-dimensional feature maps obtained from the time-frequency maps are used as inputs to train the Swin Transformer model, achieving intelligent fault diagnosis of rolling bearings. The experimental results indicate that the method combining VMD-CWT with Swin Transformer network has better fault diagnosis accuracy, with the accuracy of the test set in the measured data up to 99.79%.

Keywords:rolling bearing; variational mode decomposition; continuous wavelet transform; Swin Transformer; fault diagnosis

0 引言

随着工业设备的广泛应用,滚动轴承在各种旋转机械设备中起着不可或缺的作用,准确识别轴承故障类型并及时采取有效维修措施至关重要。

近十年来,随着深度学习领域的快速发展,以卷积神经网络(convolutional neural network, CNN)为代表的深度学习模型已被广泛应用于轴承故障诊断领域^[1]。但 CNN 并不具备全局特征捕获能力,且网络权重无法更改,无法建立长距离的特征关系^[2]。Transformer 是 GOOGLE 团队于 2017 年提出的基于纯注意力机制模型,具有优秀的长距离相关性建模能力和全局特征捕获能力^[3]。多头注意力机制(multi-head self-attention)由多层自注意力机制拼接而成,自注意

力机制可看作一种非局部滤波操作,通过估计所有位置的注意力分数并根据分数收集相应的输入来计算序列中每个位置的响应^[4]。截止目前,Transformer 的变体 ViT^[5]、CVT^[6]等模型已被成功应用于图像分类、目标检测等计算机视觉领域(computer vision, CV),但 Transformer 模型参数量大,计算成本较高,且其独特的输入格式使得模型很难直接应用于轴承故障诊断领域,但这依然阻止不了专家学者们探索的步伐:林昱涛等^[7]利用 Transformer 网络搭建轴承声-振融合模型,加强了模型时序特征提取能力,模型诊断性能有所提升;刘俊峰等^[8]针对滚动轴承变工况问题,从信号序列特性出发,将改进傅里叶分解和 Transformer 网络相结合,设计了基于注意力机制的 Transformer-CNN 序列建模单元,有效地实现了滚动轴承的故障诊断。上述方法成功将

基金项目:民航发展基金教育人才项目(MHJY2022031)

第一作者简介:曾信凌(1998—),男,重庆涪陵人,硕士研究生,研究方向为轴承故障诊断,735547510@qq.com。

Transformer 模型引入到轴承故障诊断领域,但依然存在模型计算量大的问题。

Swin Transformer 是 2021 年微软亚洲研究院提出的 Transformer 变体模型^[9]。该模型在 ViT 的基础上引入了局部注意力思想,采用移动窗口机制,不仅降低模型的计算复杂度,还提高了模型特征提取能力。因此 Swin Transformer 网络理论上更适合进行滚动轴承故障智能诊断。

变分模态分解 (variational mode decomposition, VMD) 是近年来提出的一种自适应信号处理方法,该方法通过迭代搜索最优解来确定每个分量的中心频率和带宽^[10],是一种完全非递归模型。它克服了经验模式分解 (empirical mode decomposition, EMD) 的模态混叠和端点效应,现已在故障诊断领域广泛应用。谢棕等^[11]利用 VMD 结合多尺度排列熵,实现了齿轮的故障诊断;张萍等^[12]将希尔伯特包络谱和 VMD 算法结合,实现了轴承故障特征频率的提取。刘秀丽等^[13]提出一种基于鲸鱼优化算法 (whale optimization algorithm, WOA) - VMD 能量熵的特征提取方法,完美解决了 VMD 参数难以确定的问题使用支持向量机 (support vector machine, SVM) 实现了对轴承信号的故障诊断。陈鹏等^[14]利用粒子群算法优化支持向量机,结合 VMD 实现了齿轮箱的故障诊断。廖晨等^[15]将 VMD 与广义 S 变换 (generalized S-transform, GST) 相结合,将一维振动信号转为二维时频图,结合 CNN 实现了滚动轴承的故障诊断。

综上,本文提出一种 VMD 和连续小波变换 (continuous wavelet transform, CWT) 相结合的二维时频图生成方法。针对 CNN、传统 Transformer 存在的问题,本文将 Swin-Transformer 作为诊断模型进行滚动轴承故障的智能诊断,从而减少模型训练参数的同时提高了模型的诊断精度。

1 变分模态分解

变分模态分解通过模态层数 k 、惩罚因子 α 、噪声容限 T 等参数,将信号分解为 k 个具有中心频率和有限带宽的本征模态分量 (intrinsic modal function, IMF)^[16],具体形式如下:

$$\min_{\{\mu_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (1)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mu_k = f(t) \quad (2)$$

式中: μ_k 为模态函数; ω_k 为模态分量的中心频率; $\delta(t)$ 为 Hilbert 变换时的冲击函数; “*” 为卷积运算。随后引入二次惩罚因子 α 以及 Lagrange 乘子 λ , 构建 Lagrange 增广函数:

$$L\{\mu_k(t), \omega_k(t), \lambda\} = \alpha \cdot \left(\sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right) + \left\| f(t) - \sum_{k=1}^K \mu_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_{k=1}^K \mu_k(t) \rangle \quad (3)$$

采用交替方向乘子法 (alternate direction method of multipliers, ADMM)^[17] 反复迭代求得 Lagrange 增广函数的鞍点,迭代公式如下:

$$\mu_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{\mu}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{\mu}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{\mu}_k(\omega)|^2 d\omega} \quad (5)$$

式中 d 为原始信号和各个本征模态函数 (IMF) 之和之间的误差。

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau(\hat{f}(\omega) - \sum_{k=1}^K \hat{\mu}_k^{n+1}(\omega)) \quad (6)$$

终止条件为

$$\sum_k \frac{\|\mu_k^{n+1} - \mu_k^n\|_2^2}{\|\mu_k^n\|_2^2} < \varepsilon \quad (7)$$

式中 ε 为判别精度,若满足上式条件则迭代终止。

2 Swin Transformer 网络模型

Swin Transformer 是微软亚洲研究院于 2021 年在 ViT 基础上提出的具有层级式结构的改进模型。Swin Transformer 可从小尺寸的图像块开始逐级向更深的 Transformer 层中合并相邻的图像块,从而形成一种层次化表达,如图 1 所示^[18]。相较于 ViT 的全局自注意力, Swin Transformer 所采用的小窗口自注意力大大降低了计算的复杂度,只要窗口大小固定,自注意力的计算复杂度就是固定的,因此模型的计算复杂度只会随着图像大小线性增长^[19]。

Swin Transformer 摒弃了传统 ViT 所用的全局自注意力计算方式,采用独立窗口进行自注意力计算,该模块被称为 Windows Multi-head Self-attention (W-MSA)。该方法将整图分割为多个大小相同的窗口,单独对每个窗口内部进行 Self-

attention。但 W-MSA 无法进行窗口与窗口之间的信息传递。为解决这个问题,模型在 W-MSA 模块后增加了 Shifted Windows Multi-head Self-attention (SW-MSA) 模块,用于连续自注意力层之间的窗口分区移动,从而实现每个窗口之间的信息传递交互,弥补了 W-MSA 模块无法实现窗口间信息交互的缺点,操作如图 2 所示。

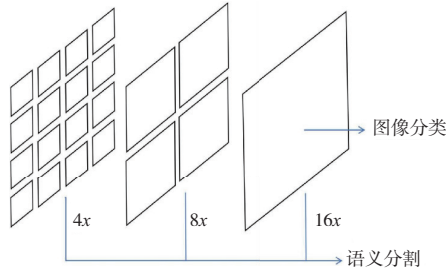


图 1 层级式结构

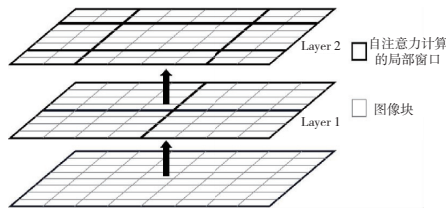


图 2 移动窗口

1 个 Swin Transformer 模块中分别包含 1 个 W-MSA 模块和 1 个 SW-MSA 模块,两种模块成对使用,如图 3 所示,这也是每个 Stage 中 Swin Transformer 模块都是偶数层的原因。Swin Transformer 模型整体结构如图 4 所示,模型图主要由分割层、层叠模块、归一化层、全局平均池化层和全连接层组成^[20]。其中 H 为图像的高度, W 为图像的宽度, C 为图像的通道数。

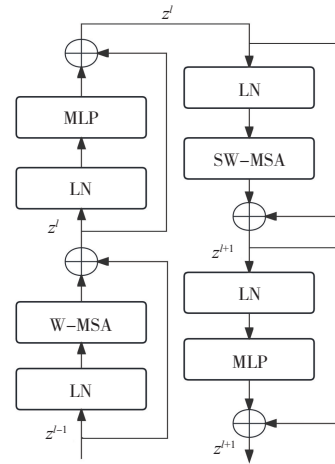


图 3 Block 的网络结构

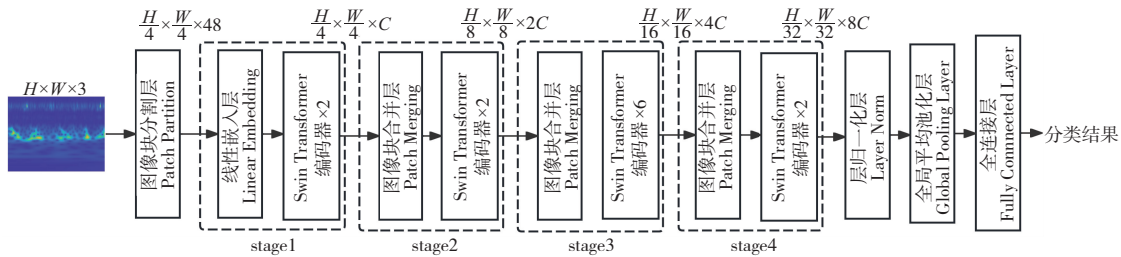


图 4 Swin Transformer 网络结构图

3 基于 VMD-CWT 和 Swin Transformer 的滚动轴承智能诊断方法

VMD-CWT 和 Swin Transformer 相结合的滚动轴承故障诊断方法如图 5 所示,具体处理步骤如下。

- 1) 使用轴承试验台采集不同故障类型的轴承振动加速度信号。
- 2) 利用鲸鱼优化算法 (WOA), 使用基尼系数作为适应度函数对 VMD 中参数 α 、参数 k 进行迭代搜索。
- 3) 通过优化后的 VMD 对一维振动信号进行

降噪处理,利用峭度指标选取合适 IMF 分量,进行信号重构。

- 4) 利用 CWT 对重构后的信号进行时频分析,并生成二维时频图像。构建基于 VMD-CWT 的二维时频图像数据,并按 7:2:1 的比例分别划分训练集、验证集、测试集。

- 5) 搭建基于 Swin Transformer 网络的诊断模型,并对模型进行初始化。

- 6) 使用训练集训练网络模型,调整网络参数以达到最优效果,并保留最优权重。

- 7) 测试集获得各类故障的分类精度,完成滚动轴承的故障诊断。

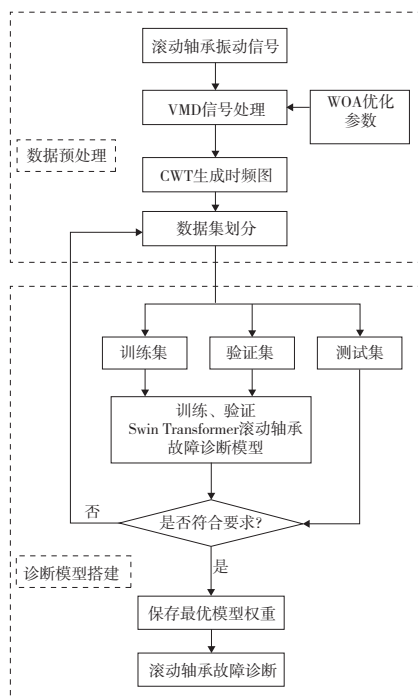


图5 VMD-CWT和Swin Transformer相结合的滚动轴承故障诊断方法流程图

4 实验验证

4.1 实验数据

为验证本文所提方法在实际应用中的有效性,采用数据集1、数据集2两个不同的数据集进行验证。前者为自制数据集,实验使用MB ER-16 K深沟球轴承,采用激光对轴承进行不同损伤类型的加工。采样频率为25.6 kHz,转速为1 500 r/min,径向载荷1 000 N。对于每种轴承状态分别取2 000个样本,5种轴承状态共10 000个样本。轴承故障信号采集实验原理如图6所示。数据集2采用美国凯斯西储大学的公开轴承数据,轴承状态包括正常、内圈故障、外圈故障和滚动体故障,每种故障100个,共400个样本。数据集1、数据集2的轴承具体故障参数如表1所示。

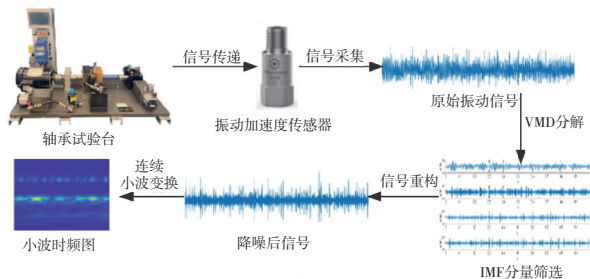


图6 实验原理图

表1 轴承故障参数

数据集编号	故障位置	故障尺寸/mm	转速/(r/min)	载荷/N
1	正常	0	1 500	1 000
1	内圈	0.325	1 500	1 000
1	外圈	0.325	1 500	1 000
1	内外圈	0.325	1 500	1 000
1	滚动体	0.500	1 500	1 000
2	正常	0	1 797	0
2	内圈	0.544	1 797	0
2	外圈	0.544	1 797	0
2	滚动体	0.544	1 797	0

4.2 信号预处理

1) VMD算法参数确定

本文使用鲸鱼优化算法对VMD参数 k 、参数 α 进行迭代寻优,并以基尼系数作为适应度函数,基于峭度指标筛选合适IMF分量,最终确定每种故障类型具体参数如表2所示。

表2 不同轴承故障VMD参数

故障类型	k_1	α_1	k_2	α_2
正常	8	2 489	8	2 631
内圈故障	5	1 090	5	1 243
外圈故障	8	3 042	5	1 348
内外圈故障	8	3 042	—	—
滚动体故障	8	1 469	7	1 441

2) 特征时频图生成

将原始故障信号进行WOA-VMD降噪后,利用CWT二维时频图生成方法对降噪后的振动信号进行处理,制作不同工况下的时频图数据集用于滚动轴承的故障诊断,数据集1的各类故障状态的二维时频图示例如图7所示。

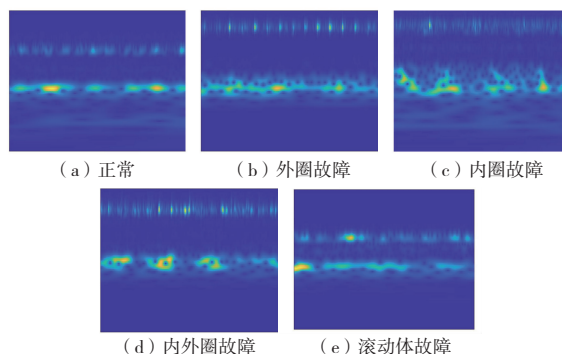


图7 不同故障的CWT时频图

4.3 Swin Transformer 网络参数

本文所用 Swin Transformer 网络模型参数如表 3 所示。

表 3 Swin Transformer 网络参数

Stage	Output Size	Swin Transformer
Stage1	56×56	Patch Size:4×4 Embed_dim:96 Window Size:7×7 depth:2 heads:3
Stage2	28×28	Patch Size:4×4 Embed_dim:192 Window Size:7×7 depth:2 heads:6
Stage3	14×14	Patch Size:4×4 Embed_dim:384 Window Size:7×7 depth:6 heads:12
Stage4	7×7	Patch Size:4×4 Embed_dim:768 Window Size:7×7 depth:2 heads:24

4.4 试验结果

本次试验配置如下:处理器为 Intel(R) Xeon(R) Gold 5218R CPU @ 2.10 GHz;显卡为 NVIDIA

RTX A6000,所用深度学习框架为 Pytorch。batch size 为 32;学习率为 1×10^{-3} ;权重衰减为 1×10^{-5} ;epoch 为 300;输入时频图大小为 224×224 ;损失函数选用交叉熵损失函数;优化器为 AdamW。

图 8 为 Swin Transformer 网络在数据集 1 下经过 300 次迭代后的准确率和损失值曲线。从图中可以看出,本研究提出的方法对不同轴承故障的时频图具有良好的分类效果。训练集的准确率在 50 次迭代后趋于 100%。同时,损失值在经过 10 次迭代后迅速下降,在迭代 100 次后趋于稳定,表明模型已经收敛。综上所述,该深度学习模型在较少的迭代次数下也能实现很高的准确率,即模型已经学习到了数据的特征,在整个训练周期内表现出良好的收敛性(本刊黑白印刷,如有疑问请咨询作者)。

为避免个别类型故障极端识别的情况,本文利用混淆矩阵和 t-SNE 非线性降维技术对测试集分类结果进行可视化分析,结果如图 9、图 10 所示。由混淆矩阵可以发现,外圈故障的分类准确率达到 100%,正常、内外圈混合故障的分类准确率为 99%,外圈和滚动体故障分类准确率为 98%。表明 Swin Transformer 网络可以有效区分易混淆的故障模型。此外,各类故障状态在聚类图中表现出明显的可分性和密集的聚簇性。进一步证明 Swin Transformer 网络优秀的性能。

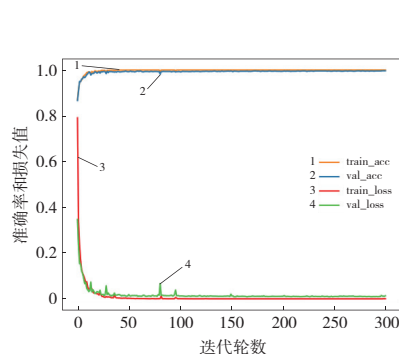


图 8 网络迭代曲线

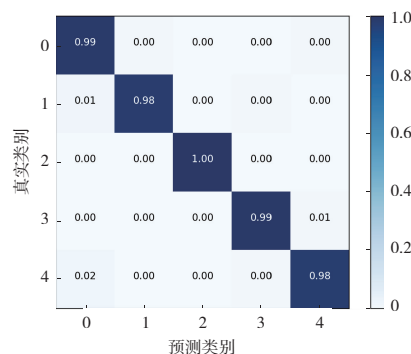


图 9 网络混淆矩阵

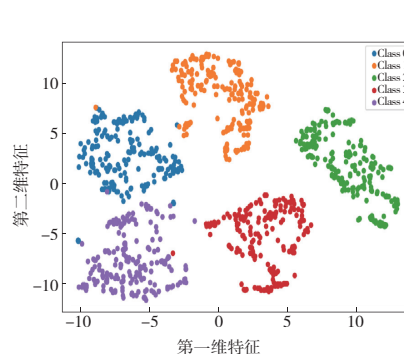


图 10 t-SNE 聚类可视化

4.5 对比分析

为进一步验证本文所提方法的优越性和 Swin Transformer 网络的泛化性。现将数据集 1、数据集 2 的一维振动信号进行 VMD 降噪,随后分别用短时傅里叶变换 (STFT)、S 变换 (ST)、广义 S 变换 (GST) 以及 VMD-CWT 方法生成各自的二维时频图作为 Swin Transformer 网络的输入,模型训练

对比结果如表 4 所示。从表中可以看出,基于 VMD-CWT 二维时频图转换结合 Swin Transformer 网络方法的准确率均高于其他时频图生成方法,说明利用 CWT 将轴承故障振动信号转化为二维图像的方法可以凸显时频域中的故障特征;Swin Transformer 模型在数据集 1 中的准确率均超过 98%,在数据集 2 中的准确率均超过 99%,在两个

不同数据集上的训练集准确率均达到 100%。这得益于 VMD 为 Swin Transformer 网络提供了深层次特征分解,CWT 突显了轴承故障信号中的多维复杂信息,使得模型能更好地提取多尺度时间频率融合特征。再加上 Swin Transformer 网络自身强大的特征提取能力,使得这种混合方法展现出卓越的性能和出色的泛化能力。

表 4 不同时频图生成方法对比

方法	训练集 1	测试集 1	训练集 2	测试集 2
VMD-CWT	100.00	99.79	100.00	99.82
VMD-STFT	98.44	97.24	99.62	99.02
VMD-ST	98.75	96.95	99.76	99.12
VMD-GST	99.45	98.34	99.82	99.43

为进一步验证 Swin Transformer 网络优越的性能,本文将 Swin Transformer 表中简称 Swin T 网络与 VGG 网络、CNN-Resnet 网络以及 ViT 进行对比。训练集准确率(数据集 1)结果如表 5 所示。

表 5 不同网络模型对比

模型	VGG	CNN	ViT	Swin T
VMD-CWT	99.44	99.34	99.69	100.00
VMD-STFT	94.83	94.63	97.84	98.44
VMD-ST	95.13	95.88	97.85	98.75
VMD-GST	96.64	95.28	98.69	99.45

从表 5 可以看出,基于时频图深度学习的滚动轴承故障诊断方法的平均识别率均在 90% 以上,将 VMD 和 CWT 相结合的轴承信号预处理方法准确率均在 99% 以上,Swin Transformer 网络更是达到了 100%,进一步证明了该方法结合 Swin Transformer 网络的优越性。同时,从表 5 中还可以看出,基于 Swin Transformer 的深度学习模型的准确率均高于传统的卷积神经网络,这是因为 Swin Transformer 的自注意力机制能够更好地捕获长距离依赖,相对于卷积层在提取局部特征的优势,Swin Transformer 能更好地捕获全局特征。综上,通过对不同模型的比较分析,基于 VMD-CWT 和 Swin Transformer 网络相结合的轴承故障诊断方法均取得最高准确率,由此证明了该方法的优越性。

5 结语

本文以滚动轴承为工程研究背景,提出一种更适合于轴承的故障诊断方法,即 VMD-CWT 结合 Swin Transformer 法。通过试验和分析可得出以下结论。

1) 该方法首先对轴承加速度寿命试验台采集到的原始振动加速度信号进行变分模态分解,寻找出合适的 IMF 进行重构降噪,随后利用连续小波变换得到二维时频图作为 Swin Transformer 网络输入,得到轴承故障诊断模型,从而实现滚动轴承各类故障的识别。

2) 本文首次将 Swin Transformer 网络应用于轴承故障诊断领域,和其他方法相比,该方法具有更好的识别精度和稳定性。在实测数据中准确率达到了 100%。所提方法可以为轴承故障诊断提供理论和技术支撑,具有较好的应用前景。

参考文献:

- [1] 陈鑫华,钱雪忠,宋威. 基于轻量级特征融合卷积网络的图像分类算法[J]. 计算机工程,2021,47(11): 268-275.
- [2] WANG H, XU J W, YAN R Q, et al. A new intelligent bearing fault diagnosis method using SDP representation and SE-CNN[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(5): 2377-2389.
- [3] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// Advances in neural information processing systems, 2017.
- [4] LIUY, ZHANG Y, WANG Y X, et al. A survey of visual transformers[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(6): 7478-7498.
- [5] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: transformers for image recognition at scale [C]// International Conference on Learning Representations, 2021.
- [6] WU H P, XIAO B, CODELLA N, et al. CvT: introducing convolutions to vision transformers [C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2021: 22-31.
- [7] 林昱涛,牛青波,马天旭,等. 基于 Transformer 的智能轴承声-振融合故障诊断[J]. 轴承, 2023(2): 67-73.
- [8] 刘俊锋,俞翔,万海波,等. 基于 MFMD 和 Transformer-CNN 的滚动轴承故障诊断方法[J]. 航空动力学报, 2023, 38(6): 1446-1456.

(下转第 34 页)

- 收稿日期:2023-11-24

40(13) : 146-153.

- 收稿日期:2023-11-14